Isolation and Sequence Analysis of ARSs and Flanking Elements from *Saccharomyces kluyveri*

Katherine Lovett

May 8th 2009

Research Advisor: Prof. Bik-Kwoon Tye
Department of Molecular Biology and Genetics
Cornell University

Abstract

DNA replication is a highly conserved process essential for the perpetuation of all life forms. Initiation of DNA replication occurs at specific sites called replication origins. The essential DNA elements that constitute these sites are not fully understood in eukaryotes. Studying origins of replication using budding yeasts as a model may provide useful information for deciphering essential components of replication origins in higher eukaryotes, including humans. Origins of replication (also known as autonomously replicating sequences or ARSs) are well characterized in the budding yeast species *Saccharomyces cerevisiae*. *S. cerevisiae* ARSs contain an ARS consensus sequence (ACS) that is essential but not sufficient for supporting replication initiation. Little is known about the essential elements of replication origins in other yeast species or in higher eukaryotes. In this study, I have cloned and analyzed the sequence of several ARSs from *Saccharomyces kluyveri*, and an ARS consensus sequence has been identified, similar to that of *S. cerevisiae* in that both ACSs are AT-rich 17 bp sequences. Additionally, I have discovered flanking elements that are essential for the function of certain *S. kluyveri* ARSs. This study provides preliminary data on the essential ACS and a flanking element that constitute functional replication origins in *S. kluyveri*.

Introduction

DNA replication is an essential and conserved process for all living organisms. During every cell division, genomic DNA must be duplicated and passed on to the new daughter cell. Errors in DNA replication lead to genetic mutations that may be deleterious to the organism. It is therefore vital for cells to carefully regulate the process of DNA replication.

Initiation is a critical step in the replication process. Eukaryotic genomes are comprised of long (typically several megabase) DNA sequences, so cells must initiate DNA replication at many points in the genome in order to accomplish replication in a timely matter. Additionally, genomes are divided into distinct units (chromosomes), and cells must initiate DNA replication on all chromosomes to ensure replication of the entire genome. However, cells must replicate their DNA once and only once during the cell division cycle, so initiation must occur no more than once at any site. Cells must therefore carefully regulate initiation to ensure complete, efficient replication without harmful redundancy (Sclafani and Holzen, 2007).

During the DNA replication process, replication is initiated from origins of replication through a series of steps involving the binding of multiple protein factors (Sclafani and Holzen, 2007). A six-subunit protein complex known as the origin recognition complex (ORC) binds to DNA at origins of replication (also known as autonomously replicating sequences or ARSs) (Bell and Stillman, 1992). ORC binds to the ARS consensus element (ACS) and recruits other protein factors to initiate replication (Nieduszynski et al., 2007). Replication then proceeds bidirectionally from the origin of replication. The six-subunit ORC is conserved across eukaryotes; however, the essential DNA elements that constitute these replication origins are not fully understood and their sequences are not conserved (Sclafani and Holzen, 2007).

Yeast are useful model organisms for studying origins of replication. Yeast, like humans, belong to the domain Eukarya. Members of this domain (eukaryotes) share certain characteristics and processes, including a similar mechanism of DNA replication (Sclafani and Holzen, 2007). Therefore, studying this process in yeast can provide information about DNA replication in other eukaryotes, including humans. Yeast are often used in research for a number of reasons. They are small and easy to maintain, and reproduce quickly, so large scale experiments are relatively inexpensive. Additionally, it is easy to manipulate yeast DNA, which makes them especially useful in studying the mechanisms of DNA replication.

In the yeast species *Saccharomyces cerevisiae*, origins of replication have been well characterized. Origins of replication have been isolated on small circular DNA molecules known as plasmids (Stinchcomb et al., 1979). Plasmids can be introduced into yeast in a process known as transformation. However, plasmids cannot be replicated if they do not have an origin of replication. Origins can be isolated or confirmed by inserting the putative origin into a plasmid and observing whether the plasmid can be propagated or maintained, because ARSs enable the plasmids to be replicated autonomously (Brewer and Fangman, 1987).

In *S. cerevisiae* approximately 325 ARSs have been identified and confirmed as origins of replication (Nieduszynski et al., 2007). Among these origins a common motif (ACS), has been identified. This ACS is A-T rich, which may facilitate unwinding of the DNA. Many of these ARSs have also been shown to be binding sites for ORC, confirming that they are origins of replication (Wyrick et al., 2001). Additionally, the ACS is flanked by "B elements" that are not as highly conserved in sequence. The ARSs, including the B elements, are typically ~200 bp in length (Nieduszynski et al., 2007).

There are at least 3 classes of B elements, which differ in function. The B1 elements interact with ORC (Lee and Bell, 1997). The B2 elements have been proposed to contain a DNA unwinding element (Lin and Kowalski, 1997), although other research has suggested they are required for loading of MCM proteins, which interact with ORC during initiation (Wilmes and Bell, 2002). The B3 elements interact with ABF1, a transcription factor that binds a subset of origins (Marahrens and Stillman, 1992). At least one of the three B elements, in combination with the ACS, is essential for ARS function. However, though these B elements have been fairly well studied in *S. cerevisiae*, little is known regarding auxiliary elements in other species.

In other eukaryotes, including the fission yeast species *Schizosaccharomyces pombe*, origins do not appear to be so well defined. There is little consensus in the *S. pombe* ARSs that have been isolated, although they are similarly A-T rich (Dai et al., 2005). Additionally, certain *S. cerevisiae* ARSs have been shown to be highly efficient, meaning they are utilized in almost every cell cycle: 72-88% in "high frequency origins", and 40-58% in "intermediate frequency orgins" (Yamashita et al., 1997). In *S. pombe*, origins generally have a much lower efficiency (~30%), with different origins being used each cell cycle (Patel et al., 2006).

In this study, I am isolating ARSs in the yeast species *Saccharomyces kluyveri* and comparing them to those of *S. cerevisiae*. *S. kluyveri* is more closely related to *S. cerevisiae* than *S. pombe* (Washington University Genome Sequencing Center, Figure 5), so the initiation process of *S. kluyveri* may more closely resemble that of *S. cerevisiae*. By comparing ARSs from these two species, common elements may be found, giving a broader picture of the essential DNA elements that constitute an origin of replication. A primary goal of this study is to identify an ACS for *S. kluyveri*. This requires isolating a large number of ARSs for sequence

analysis, and several student researchers have been involved in isolating *S. kluyveri* ARSs in a class project, which is the precursor of this project (see Acknowledgments).

Once isolated, ARSs can be truncated to find the minimum sequence required for ARS function. They can also be mutagenized to identify critical base pairs in the sequence. These data can then be analyzed to determine a consensus sequence for *S. kluyveri*. Additionally, *S. kluyveri* ARSs can be tested for ARS function in *S. cerevisiae* and vice versa, to gain more information about the essential components of replication origins as well as the replication initiation machineries in both species. In this study I have used truncation and mutagenesis to analyze certain *S. kluyveri* ARSs. In addition, I identified several ARS flanking elements, which may be equivalent to the B elements of *S. cerevisiae*. Altogether, this work broadens our understanding of origins of replication in yeast, which may improve our understanding of initiation of DNA replication in all eukaryotes.

Experimental Procedures

Bacterial Transformation

Bacterial transformations were conducted using a modified version of the Bioline protocol (Bioline, n.d.). Bioline α-select chemically competent cells were thawed on ice, and 20 μL of cells were transferred to a microfuge tube. The DNA solution was then added (1 μL of mini-prepped plasmid, 2 μL of cloning reaction, or 10 μL of plasmid obtained from yeast following the plasmid pullout protocol. DNA concentration varied depending on the source, and the number of colonies therefore varied, ranging from approximately 50-400.). The mixture was incubated on ice for 30 min, incubated at 42°C for 30 seconds to induce heat shock, and replaced on ice for 2 minutes. 150 μL of LB was then added and the cells were incubated at 37°C for 1

hour before being plated on LB + ampicillin plates. Transformants were grown overnight at 37°C.

Bacterial Mini-prep

Bacterial mini-preps were conducted following the Promega Wizard Plus SV Mini-Prep protocol (Promega, n.d.), using 2mL of cell culture and eluting the DNA in 50 µL of water instead of 100 µL.

Yeast Transformation

Yeast transformations were conducted following the protocol designed by Ivan Liachko (personal correspondence). Cells were grown overnight in liquid culture at 30°C, and 1mL of culture was centrifuged in a microfuge tube. Supernatant was removed, and DNA was added (1 µL of mini-prepped plasmid or 5 µL of genomic library. DNA concentration varied depending on the source, and the number of transformants therefore varied, ranging from approximately 10-200). 75 µL of 1-step buffer (0.2M Lithium Acetate, 40% PEG 4000, 100mM DTT, 0.5 mg/mL carrier DNA) was added and the solution was vortexed to resuspend the cells. Cells were incubated at 42°C for 1 hour, plated on selective media, and grown at 30°C for 3-5 days.

Plasmid Recovery from *S. kluyveri*

Plasmids were recovered from *S. kluyveri* following the protocol designed by Ivan Liachko and Justin Donato (personal correspondence). 1 mL of saturated yeast culture was centrifuged in a microfuge tube, and the supernatant was removed. Cells were resuspended in 500 µL of Y1 buffer (1M sorbitol, 0.1M EDTA, pH 7.4, 1 µL/mL fresh β-mercaptoethanol). 50 µL of 5mg/mL yeast lytic enzyme were added, and the solution was incubated at 30°C for 1 hour. The solution was then centrifuged and the supernatant was removed. The bacterial mini-prep procedure was then used to obtain the plasmids.

Construction of pIL17

To screen for more flanking elements, a new plasmid (pIL17) was constructed using the plasmid pRS406 (Table 2), which contains the *URA3* gene and a BamHI site. Primers (Table 3) were designed to amplify EC175min (an *S. kluyveri* ARS truncated by Ivan Liachko to 91 bp) and add a BglII restriction site on one side of the ARS. The primers also added an MboI site to the ends of the sequence. The ARS with the BglII site was cloned into pRS406 at the BamHI site, generating the new plasmid pIL17.

Preparation of genomic DNA

FM628 cells were broken by vortexing with glass beads in lysis buffer (17% glycerol, 50mM MOPS, 150mM potassium acetate, 2mM magnesium chloride, 500uM spermidine, and 150uM spermine; pH7.2). Spheroplasts were collected by centrifugation for 10 minutes at 8000 rpm (4C) and resuspended in G2 buffer (Qiagen). RNase A (Qiagen) and proteinase K (Invitrogen) were added to the buffer to 200ug/ml and 400ug/ml final concentration respectively and the solution was incubated at 37°C for 4 hours with gentle shaking every 30 minutes. The solution was centrifuged for 5 minutes at 5000 rpm (4°C) and the supernatant was passed through Genomic-Tip 100/G column (Qiagen) to purify DNA. Wash and elution was performed according to Qiagen Genomic-Tip 100/G manual. DNA was precipitated from the eluate with isopropanol and resuspended in 500ul distilled water.

Generation of genomic libraries

To generate a genomic library, genomic DNA from the *S. kluyveri* strain FM628 (Table 1) was digested with a restriction enzyme and then cloned into a plasmid. For the plasmid pIL07 (Table 2), genomic DNA was digested with MboI and cloned into pIL07 at the BamHI restriction site. For the plasmid pIL17, genomic DNA was digested with AluI, RsaI, or HaeIII, treated with

antartic phosphatase, and then cloned into pIL17 at the BglII restriction site. After cloning, the collection of plasmids was transformed into *E. coli* (2 μL cloning reaction per bacterial transformation) and the cells were grown overnight. The transformants were collected by adding 2.25 mL water to the plate and scraping with a glass spreader. The suspension of cells was transferred to a microfuge tube and mini-prepped to generate a genomic library containing multiple copies of each insert-containing plasmid.

Isolation of ARSs and ARS-enabling flanking elements

To isolate new ARSs and ARS-enabling flanking elements, the genomic library was transformed (see above) into FM628 (Table 1), and the transformants were plated on CSM-leu-ura media (Sunrise Science Products, complete supplemental mixture minus leucine and uracil) and grown for 3-5 days at 30°C to select for those that contained a maintainable plasmid. Colonies growing on the CSM-leu-ura media were restreaked on a new CSM-leu-ura plate to retest. Only one colony was selected from each plate of transformants generated from a single library to decrease the chances of isolating the same ARS multiple times. Cultures were inoculated from these strains in CSM-leu-ura broth, and were grown for 2 days at 30°C. The plasmid was then obtained from the cultures following the plasmid recovery protocol (see above). The plasmid was then transformed into *E. coli* (see above) to produce more copies of the plasmid. The plasmid was then was transformed back into FM628, plated on CSM-leu-ura, and grown at 30°C for 3-5 days to confirm that it contained a functional ARS. Once the ARS was confirmed, the insert was sequenced by the Cornell University Life Sciences Core Laboratories Center, using primers specific to the plasmid (IL325 and IL336 for pIL07, or IL429 and IL430 for pIL17). ARS sequences were analyzed using a motif search algorithm developed by Dr. Uri Keich to identify the consensus sequence.

In order to determine whether an insert isolated in pIL17 is an ARS or a flanking element, it must be tested for ARS function in pRS406. Primers were designed to amplify the insert from pIL17 and add MboI sites to the ends. This insert can then be digested with MboI, cloned into pRS406 at the BamHI site, and transformed into FM628 to test for ARS function.

Characterizing ARSs: truncation

To localize the ARS on cloned inserts, I carried out truncation analysis on three of the isolated ARSs. The ends of the ARSs were arbitrarily designated "left" or "right" depending on their orientation in the plasmid. Primers were designed such that a portion of the insert could be amplified from the plasmid using PCR, generating a truncated version of the insert (Table 3). The primers also added an MboI site to the ends of the truncated sequence. The truncated inserts were digested with MboI and cloned into pIL07 (Table 2) at the BamHI restriction site. The new plasmids, containing the truncated inserts were then transformed into FM628 and tested for ARS function. Truncation schemes varied depending on the length of the original isolated ARS (Figure 3).

Characterizing ARSs: site directed mutagenesis

To determine the critical bases in the ARS, site-directed mutagenesis using site-specific primers was performed on the truncated ARS KL4min. Primers were designed such that they were perfectly homologous to the ARS except for a 3 bp sequence (Table 3). PCR was used to amplify the insert with the 3 bp change, producing a mutated version of the insert. Our mutagenesis strategy was usually directed at a stretch of 3 Ts, converting them to Gs. The primers also added an MboI site to the ends of the truncated sequence. The mutated inserts were digested with MboI and cloned into pIL07 (Table 2) at the BamHI site. The new plasmids, containing the mutated inserts were then transformed into FM628 and tested for ARS function.

Isolation of a flanking element from pIL07

Because the truncated ARSs only function in one orientation in pIL07, there must be some element on the plasmid that either enables the ARS to function (the ARS can only function when this element is present) or prevents the ARS from functioning (the ARS cannot function when this element is present). To determine whether the plasmid pIL07 was enabling or preventing ARS function, the truncated ARSs were cloned into the plasmid pRS406, which contains the *URA3* gene and a BamHI site. The truncated ARSs were digested with MboI and then cloned into pRS406 at the BamHI site. The plasmids with the inserts were then transformed into FM628 and tested for ARS function. As a control, the untruncated ARSs (KL4 and EC175) were cloned into pRS406 and tested for ARS function.

To determine what portion of pIL07 was enabling ARS function in the truncated ARSs, primers were designed such that they complemented the plasmid 100 bp from the site of the insert, on either side. PCR was used to amplify the insert plus a small portion of pIL07 as a single unit (Figure 4). The primers also added an MboI site to the ends of the sequence. These new "extended ARSs" (containing the truncated ARS plus a small portion of pIL07) were then cloned into pRS406 at the BamHI site and transformed into FM628 to test for ARS function.

Results

Identification of ACS of *S. kluyveri*

To isolate new ARSs, random fragments of *S. kluyveri* genomic DNA were cloned into plasmids, generating genomic libraries. The libraries were then transformed into *S. kluyveri*, and the transformants were plated on appropriate media to select for those that retained the plasmid. The backbone plasmid cannot be maintained because this plasmid lacks an ARS, so colonies are unable to form. If the insert contained an ARS, the plasmid could then be replicated and

maintained and a colony would form. Plasmids that enabled growth on the selective media were isolated from the yeast and retested to confirm ARS function before sequencing.

A total of 120 ARSs have been isolated from *S. kluyveri* using pIL07, through the efforts of several researchers (Ivan Liachko and Claire Chung isolated the majority of the ARSs [105]. Also, the BIOBM 399 class of Spring 2008 [under the guidance of Maki Inada, Ivan Liachko and Claire Chung] isolated 15 new ARSs. This class consisted of: myself, Lu Yang, Heather Pace, Eugene Cha, Lindsay Hallas, Arael Candelaresi, Zubair Azad, Kinsha Baidya, and Gina Kang.). These ARSs have been analyzed using a motif search algorithm developed by Dr. Uri Keich to identify the consensus sequence (Figure 1). It appears there is indeed a consensus sequence in *S. kluyveri*, and it is a 17 bp sequence that is T rich on one strand and has a strong resemblance to the ACS of *S. cerevisiae* (Figure 1).

Localization of ARSs on cloned inserts by truncation

To localize the ARS on cloned inserts, I carried out truncation analysis on three of the isolated ARSs: HP171, AC16, and KL4 (Figure 2). Inserts were typically a few hundred base pairs long, but past truncation experiments have shown that the ARSs from *S, kluyveri* can be truncated down to a ~100 bp sequence (Ivan Liachko, personal correspondence), that is sufficiently short so as to make mutagenesis analysis convenient. HP171, AC16, and KL4 were 624 bp, 776 bp, and 273 bp respectively. The ends of the ARSs were arbitrarily designated "left" or "right" depending on their orientation in the plasmid.

During the course of this study I partially truncated the ARS HP171, which was isolated as a 624 base pair sequence (Figure 2). I removed 250 base pairs from the right or left side of HP171, leaving 374 base pair sequences (HP171-250R and HP171-250L, respectively). I found that removing the right side of the ARS (HP171-250R) resulted in a functional ARS, while

removing the left side (HP171-250L) resulted in lack of function. These results indicate that the functional domain (HP171min) is located somewhere in the left 374 base pairs of HP171.

I partially truncated the ARS AC16, which was isolated as a 776 base pair sequence (Figure 2). I removed 300 base pairs from the right or left side of AC16, leaving 476 base pair sequences (AC16-300R and AC16-300L, respectively). Based on colony formation in the transformation assay, I found that removing the right side of the ARS (AC16-300R) resulted in a functional ARS, while removing the left side (AC16-300L) resulted in lack of function. These results indicate that the functional domain (AC16min) is located somewhere in the left 476 base pairs of AC16.

I truncated the ARS KL4 down to an 89 base pair functional domain, KL4min. (Figure 2). KL4 was isolated as a 273 base pair sequence. I removed 100 base pairs from either side of KL4, leaving a 173 base pair sequence. I found that removing the right side of the ARS resulted in a functional ARS, and removing the left side also resulted in a functional ARS. These results indicate that the functional domain (KL4min) is located in the 73 base pairs in the middle of KL4. *S. kluyveri* ARSs had previously been truncated to ~100 bp sequences (Ivan Liachko, personal correspondence, see above) so for KL4 the middle 89 base pairs were designated KL4min. This truncated ARS was found to be functional, and no further truncations were carried out.

Verification of ACS by directed mutagenesis

After finding the functional domain KL4min, I used site-directed mutagenesis to determine which base pairs were critical for ARS function (Figure 3). Based on the T-rich ACS identified by motif search (Figure 1), T stretches were targeted and replaced with Gs, because Gs and Cs are uncommon so they seemed most likely to disrupt ARS function. The mutagenized

ARSs were then tested for ARS function. I generated 7 mutant ARSs, with mutations spaced roughly equally throughout KL4min. For 6 out of 7 of the mutations, the ARS was still able to function. For one mutation, the ARS became nonfunctional. This mutation was approximately in the middle of the KL4min, in a 3 base pair stretch of A-Ts that was near a 12 base pair stretch of A-Ts. Based on the ACS identified by motif search and the localization of the ACS by site directed mutagenesis, the best match for the ACS in KL4min is **TTTTTTATGTTTCCTAT.**

Isolation of an ARS-enabling flanking element

During the truncation process, we found that certain truncated ARSs (KL4 and EC175, an *S. kluyveri* ARS truncated by Ivan Liachko to 91 bp) were functional when oriented in one direction within the plasmid pIL07 but were not functional when they were oriented in the opposite direction (based on colony formation in the transformation assay). During the cloning process, inserts have a roughly equal chance of being cloned into the plasmid in either orientation. A single cut is made in the plasmid, and the insert is identical on either end, so it can be cloned into the gap in either direction. Orientation was not expected to affect plasmid function, however. The fact that the truncated ARSs were unable to function in one orientation suggested that there was some element on the plasmid that was either enabling or preventing ARS function in a particular orientation. Previous research has shown that the surrounding DNA sequences can affect origin function (Nieduszynski et al., 2005).

To determine whether the plasmid pIL07 was enabling or preventing ARS function, the truncated ARSs were cloned into a new plasmid, pRS406. pRS406 lacks an ARS and contains the *URA3* gene, so the same assay can be used for this plasmid as was used for pIL07. The plasmids were transformed into FM628 and tested for ARS activity. They were both found to be non-functional, suggesting that there is an element on pIL07 that was enabling them to function.

As a control, the full-length ARSs fragments were cloned into pRS406. The resulting plasmids were transformed into FM628 and KL4 was found to be functional, showing that cloning the ARS into pRS406 does not prevent ARS function. EC175 did not function in pRS406, indicating that the full-length EC175 (273 bp) may also be dependent on the flanking element for function.

To determine what portion of pIL07 was enabling ARS function in the truncated ARSs, "extended ARSs" (containing the truncated ARS plus a small portion of pIL07) were cloned into pRS406 and tested for ARS function. KL4min/URA100 and EC175min/URA100 contained KL4min or EC175min respectively plus100 bp of the URA3 side of pIL07 (Figure 4). Likewise, KL4min/LEU100 and EC175min/LEU100 contained KL4min or EC175min respectively plus 100 bp of the LEU2 side of pIL07 (Figure 4). I found that both KL4/minURA100 and EC175min/URA100 were functional ARSs (based on colony formation in the transformation assay), while KL4min/LEU100 and EC175min/LEU100 were not functional ARSs (Table 4). It was concluded that there is some "flanking element" in pIL07, 100 bp or fewer in size, on the *URA3* side of the BamHI site, that is not itself an ARS but enables the truncated ARSs to function.

Isolation of new ARSs and ARS-enabling flanking elements

To screen for more ARS-enabling flanking elements, I constructed a new plasmid (pIL17) by inserting EC175min into the plasmid pRS406. EC175min alone does not function as an ARS but in the presence of the flanking element found in pIL07 it gains ARS function. By inserting EC175 min into the plasmid pRS406 (which does not contain the flanking element) and then cloning random genomic DNA fragments into the plasmid, I can screen for sequences that function like the flanking element found in pIL07 and enable EC175min to function as an ARS.

In a process similar to the initial ARS screen, I generated a genomic library by cloning random *S. kluyveri* genomic DNA into pIL17. This library was then transformed into FM628, which lacks the *URA3* gene, and the transformants were plated on CSM-leu-ura media to select for those that contained a plasmid with a functional ARS. If the plasmid contained a functional ARS, it could then be replicated and maintained and a colony would form. Plasmids that enabled growth on the selective media were isolated from the yeast and retested to confirm ARS function before sequencing.

From this screen alone, it is impossible to determine whether the new inserts isolated are flanking elements or if they are ARSs themselves. Either would enable the plasmid to be maintained. In order to determine whether the isolate is itself an ARS, it was tested for ARS function in pRS406. If the insert is not itself a functional ARS, but it enables pIL17 to be maintained, it is considered a flanking element because it enables the function of the truncated ARS EC175min.

To date a total of 60 putative ARSs or flanking elements have been isolated using pIL17 (Ivan Liachko isolated 18, and I isolated 42). Eighteen of them have been analyzed to determine whether they are ARSs or flanking regions (Ivan Liachko, personal correspondence). Eleven of them are ARSs, some of which had never been isolated using pIL07. These new ARSs can be analyzed along with those isolated using pIL07 to obtain a more accurate ACS. Correspondingly, seven of the new isolates contain ARS-enabling flanking elements. There is not yet enough data to determine a consensus sequence among the flanking elements. We are in the process of investigating whether these flanking elements are located near any of the cloned ARSs in their native genomic locations.

Discussion

An important finding of this study is that *S. kluyveri* ARSs are closer in size to those of *S. cerevisiae* than to those of *S. pombe* and other eukaryotes, and from current analysis *S. kluyveri* ARSs appear to have an ACS similar to that of *S. cerevisiae*. Both appear to be A-T rich, with an ATG in the middle, although in *S. kluyveri* TTG is slightly more common than ATG, while in *S. cerevisiae* ATG is much more common. In addition, the ACS of *S. kluyveri* appears to be more T-rich on a single strand than the *S. cerevisiae* ACS. Sequences that are T-rich on one strand and A-rich on the other bend DNA in a structure that may facilitate DNA melting or ORC recognition. Further study is required to determine the mechanism by which the T-rich ACS facilitates initiation, but this study shows that the mechanism is probably conserved across the species because for both *S. cerevisiae* and *S. kluyveri* the ACS is T-rich on one strand and A-rich on the other. The ACS for *S. kluyveri* can be refined as more ARSs are isolated, providing a larger sample size. Truncation analysis can also contribute to a more accurate ACS by removing extraneous sequences.

Mutagenesis, which can determine the portions of the ARS that are necessary for function, can also contribute to determining the ACS. Mutating ARSs effectively provides a greater stock of ARSs to analyze. Mutations that cause the ARS to become nonfunctional are likely to be in bases that are conserved across the ARSs and are part of the ACS. Examining the location of the mutation that caused KL4min to become nonfunctional reveals that it is in a region with a sequence (...**TTTTTTATGTTTCCTAT...**) similar to the ACS (Figures 3 and 1). When the 3 Ts following the ATG are converted to Gs (...**TTTTTTATGGGGCCTAT...**), the ARS becomes nonfunctional. This suggests that this mutation has interrupted the ACS and thus prevents the ARS from functioning. The truncation process removed most possible matches to

the ACS, making this sequence a highly likely candidate. Additionally, the lack of function caused by the mutation is not orientation dependent, suggesting that it is not dependent on the plasmid context. Together this data strongly suggests that the mutation interrupted the ACS, causing the ARS to become nonfunctional. The ATG in this sequence is probably the conserved ATG or TTG identified in the *S. kluyveri* ACS, so the ACS for KL4min probably is **TTTTTTATGTTTCCTAT.** Replacing the Ts with Gs may be inhibiting DNA melting or it may have disrupted the ORC recognition site.

An exciting result of this study was the discovery of the flanking element in pIL07 that enabled the ACSs to function. Previous research had shown that *S. cerevisiae* ACSs are often flanked by B elements that enhance ARS function and are conserved in function but poorly conserved in sequence (Nieduszynski et al., 2007). However, flanking sequences have not been documented in other species, including *S. kluyveri*. Discovering the existence of such a sequence opened up a new avenue of research in this species. Early research has shown that other flanking elements can be isolated which are not ARSs themselves but enable truncated ARSs to function. As research continues, we will be testing whether these elements are similar to the *S. cerevisiae* B elements.

In the broader scheme, this study provides information regarding origin recognition in yeast in general. Previously *S. cerevisiae* was the only known eukaryote to have a well conserved consensus sequence. The ARSs from the yeast species (*S. pombe*) were much larger and were poorly conserved (Dai et al., 2005), as were those of other eukaryotes (Sclafani and Holzen, 2007). This study begins to show that a conserved ACS such as that of *S. cerevisiae* is not unique to the species but can be found in other species. In addition, it shows that other species may have ARS flanking elements similar in function to those of *S. cerevisiae*. As more

data is collected concerning *S. kluyveri* and other yeast species, a more complete picture of origin recognition and initiation of DNA replication can be formed. Understanding the process of origin recognition in yeast may then yield a greater understanding of this process in eukaryotes in general, including humans. Additionally, it can provide an understanding of the evolution of the process of origin recognition and the initiation of DNA replication. *S. kluyveri* is more closely related to *S. cerevisiae* than either are to *S. pombe*, (Figure 5) and this work can help bridge the gap and reveal information regarding the evolutionary changes between the species.

Immediate future research efforts should continue to focus on *S. kluyveri* ARSs. Isolation of additional ARSs, truncation, and mutagenesis can improve the identity of the consensus sequence. As a more complete picture of ARS distribution in the *S. kluyveri* genome emerges, studies could be done concerning the efficiency of these ARSs. Are they highly efficient like the *S. cerevisiae* ARSs, used in most cell cycles? Or are they more like the *S. pombe* ARSs, which are apparently used stochastically and with low efficiency? Additionally, testing *S. kluyveri* ARSs for function in *S. cerevisiae* and vice versa can provide information concerning how the process is conserved across species.

The flanking element should also be studied in much greater depth, by isolating these elements and looking for a consensus sequence, as well as comparing the effectiveness of different flanking elements on different truncated ARSs. Dependence on flanking elements has only been observed for KL4min and EC175min, but perhaps if other ARSs were further truncated, they too would require a flanking element to function. Additional truncations of isolated ARSs are therefore another path to increasing our understanding of the flanking elements.

Further research could focus on other yeast species, including isolating ARSs, searching for an ACS, and comparing the function of ARSs from various species across evolution. Such studies may provide insights into the mechanism of origin recognition in yeast. They may also provide an avenue for the study of the co-evolution of the origins of replication and the machineries involved in origin recognition and DNA initiation through cross-species functional analyses. Similar approaches may also apply to the study of origin recognition and initiation of DNA replication in other organisms.

References

Bell, S.P., and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature *357*, 128-134

Bioline. Product Insert α-Select Chemically Competent Cells. Retrieved May 5, 2009, from http://www.bioline.com/documents/product_inserts/Alpha-Select%20Gold%20Efficiency.pdf

Brewer, B.J., and Fangman, W.L. (1987). The localization of replication origins on ARS plasmids in S. cerevisiae. Cell *51*, 463–71

Cliften, P., Sudarsanam, P., Desikan, A., Fulton, L., Fulton, B., Majors, J., Waterston, R., Cohen, B.A., Johnston, M. (2003) Finding functional features in *Saccharomyces* Genomes by phylogenetic footprinting. Science *301*, 71-76 Supplemental figure *Saccharomyces* Phylogeny. Retrieved May 8, 2009 from http://www.genetics.wustl.edu/saccharomycesgenomes/yeast_phylogeny.html

Dai, J., Chuang, R.Y., and Kelly, T.J. (2005). DNA replication origins in the Schizosaccharomyces pombe genome. Proc. Natl. Acad. Sci. USA *102*, 337–42

Lee, D.G. and Bell, S.P. (1997). Architecture of the yeast origin recognition complex bound to origins of DNA replication. Mol. Cell. Biol. *17*, 7159–7168

Lin, S., and Kowalski, D. (1997). Functional equivalency and diversity of *cis*-acting elements among yeast replication origins. Mol. Cell. Biol. *17*, 5473-5484

Marahrens, Y. and Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. Science *255*, 817–823

Nieduszynski, C.A., Blow, J.J., and Donaldson, A.D. (2005). The requirement of yeast replication origins for pre-replication complex proteins is modulated by transcription. Nucleic Acids Res. *33*, 2410-2420.

Nieduszynski, C.A., Hiraga, S., Ak, P., Benham, C.J., and Donaldson, A.D. (2007). OriDB: a DNA replication origin database. Nucleic Acids Res. *35*, D40–46

Patel, P.K., Arcangioli, B., Baker, S.P., Bensimon, A., and Rhind, N. (2006). DNA replication origins fire stochastically in fission yeast. Mol. Biol. Cell *17*, 308–16

Promega. Technical Bulletin Wizard® Plus SV Minipreps DNA Purification System. Retrieved May 5, 2009, from http://www.promega.com/tbs/tb225/tb225.pdf

Qiagen. Genomic DNA Handbook – English. Retrieved May 7, 2009, from http://www1.qiagen.com/Products/GenomicDnaStabilizationPurification/QiagenGenomicTipSystem/QIAGENGenomic-tip100G.aspx#Tabs=t2

Sclafani, R.A., and Holzen, T.M. (2007). Cell Cycle Regulation of DNA Replication. Annu Rev Genet. *41*, 237-280

Stinchcomb, D.T., Struhl, K., and Davis, R.W. (1979). Isolation and characterization of a yeast chromosomal replicator. Nature *282*, 39-43.

Wilmes, G.M. and Bell, S.P. (2002). The B2 element of the *Saccharomyces cerevisiae* ARS1 origin of replication requires specific sequences to facilitate pre-RC formation. Proc. Natl Acad. Sci. USA *99*, 101–106

Wyrick, J.J., Aparicio, J.G., Chen, T., Barnett, J.D., Jennings, E.G., Young, R.A., Bell, S.P., and Aparicio, O.M. (2001). Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins. Science *294*, 2357–60

Yamashita, M., Hori, Y., Shinomiya, T., Obuse, C., Tsurimoto, T., Yoshikawa, H., and Shiahige, K. (1997). The efficiency and timing of initiation of replication of multiple replicons of Saccharomyces cerevisiae chromosome VI. Genes Cells *2*, 655-665

Acknowledgements

I would like to thank Dr. Bik-Kwoon Tye for her guidance and feedback in writing this thesis and for trusting me to work in her lab. I would like to thank Ivan Liachko for guiding me through this project, answering my incessant questions, and enabling me to gain competency and confidence in the lab. I would also like to thank Chanmi Lee for helping me out whenever I needed it, especially in finding everything I needed to run my experiments. Also Dr. Maki Inada, for getting me started in research, and for continued guidance and encouragement.

Additionally I would like to acknowledge the work of the other researchers who helped isolate *S. kluyveri* ARSs and contributed data to the project. Ivan Liachko and Claire Chung isolated the majority of the ARSs (105). Also, the BIOBM 399 class of Spring 2008 (under the guidance of Maki Inada, Ivan Liachko and Claire Chung) isolated 15 new ARSs. This class consisted of: myself, Lu Yang, Heather Pace, Eugene Cha, Lindsay Hallas, Arael Candelaresi, Zubair Azad, Kinsha Baidya, and Gina Kang.

Figure 1

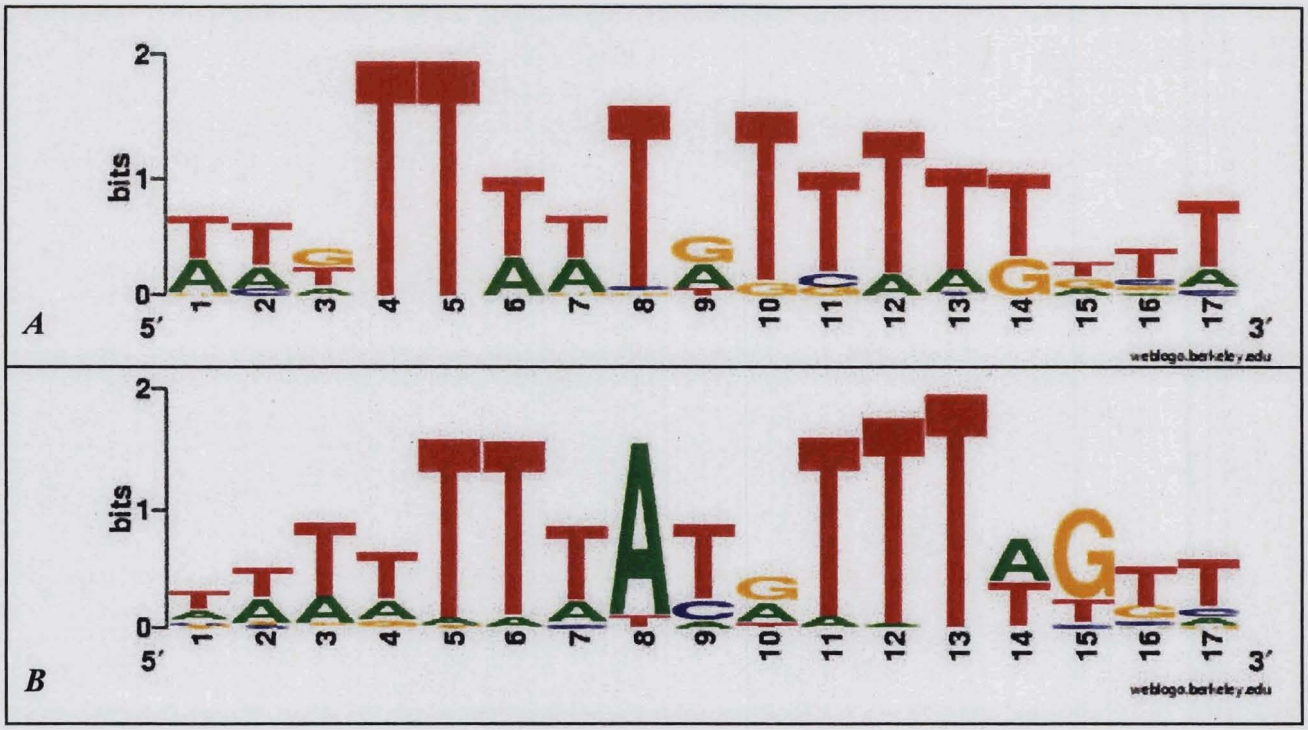

Figure 1: Consensus sequences for *S. kluyveri* and *S. cerevisiae*, analyzed using a motif search algorithm developed by Dr. Uri Keich. The height of the letter corresponds to the frequency of the occurance of the base at the location. **(A)** Preliminary consensus sequence for *S. kluyveri*, derived from 120 isolated ARSs. **(B)** Consensus sequence for *S. cerevisiae*, derived from 39 ARSs.

Figure 2

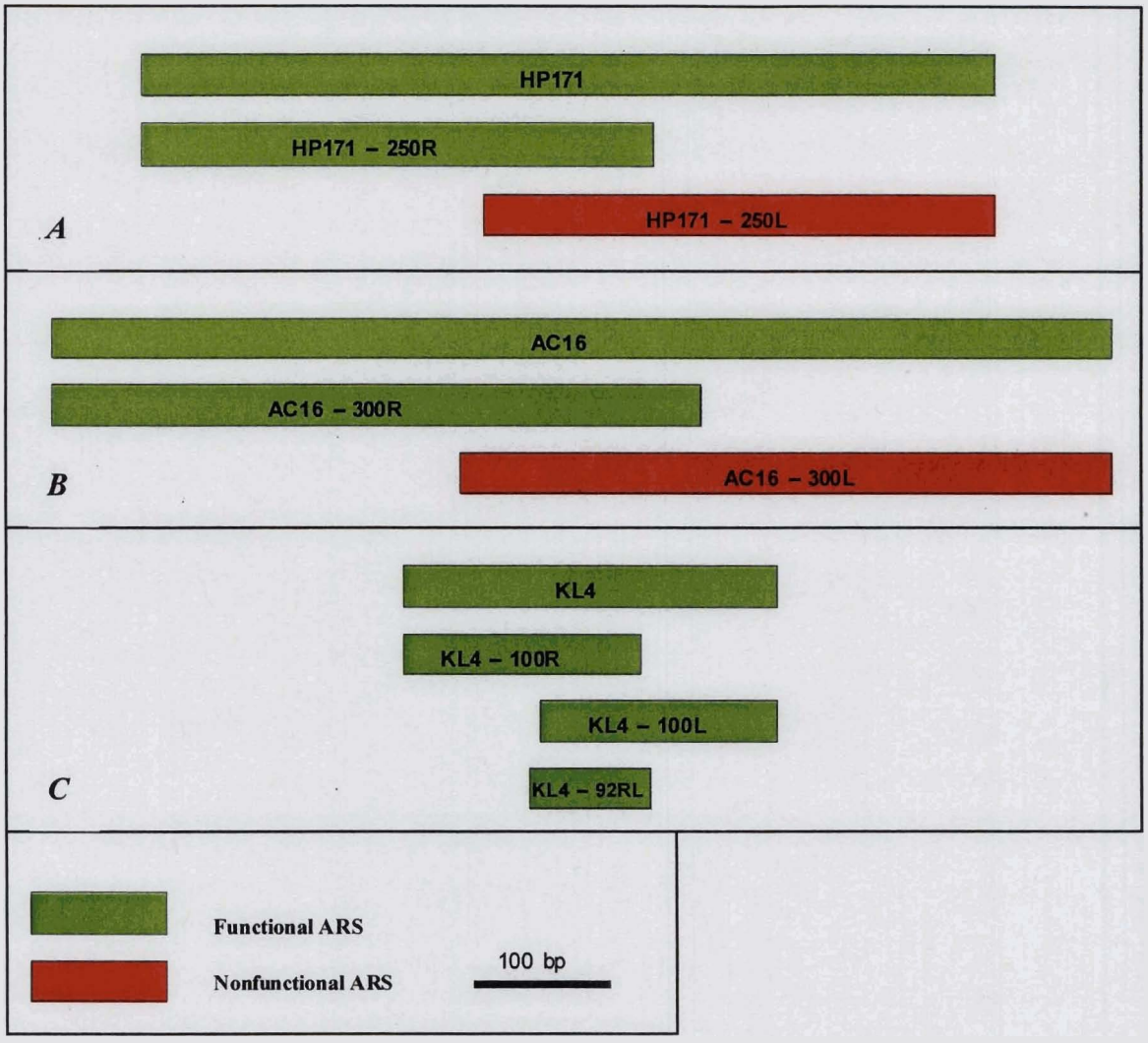

Figure 2: Diagram of truncations showing functionality of truncated *S. kluyveri* ARSs. ARSs were truncated by removing bases from the right (-XR) or from the left (-XL) **(A)** Truncations of HP171 **(B)** Truncation of AC16 **(C)** Truncations of KL4

Figure 3

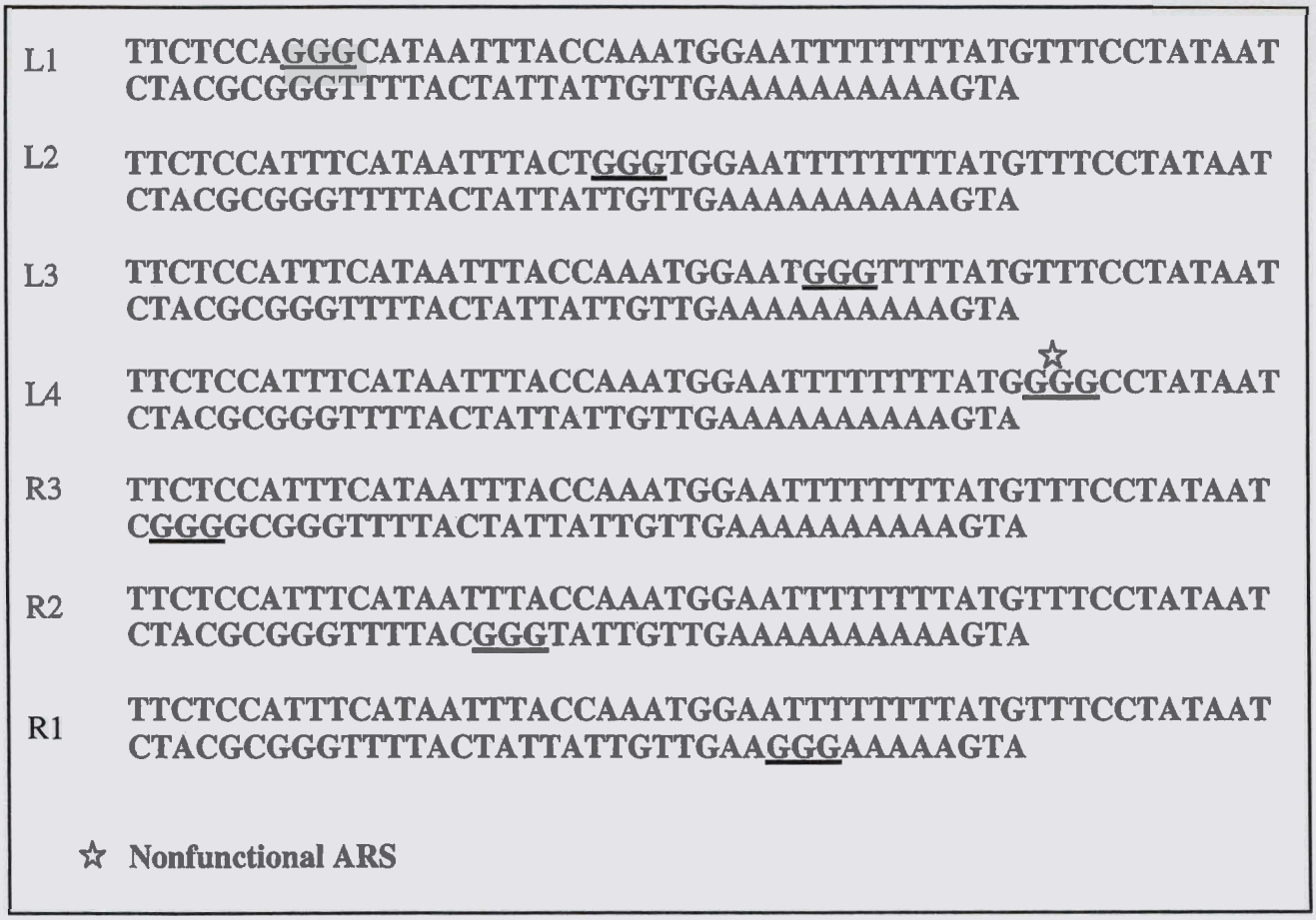

Figure 3: Site-directed mutagenesis of KL4min, an ARS isolated from *S. kluyveri* and truncated to an 89 bp sequence.

Figure 5

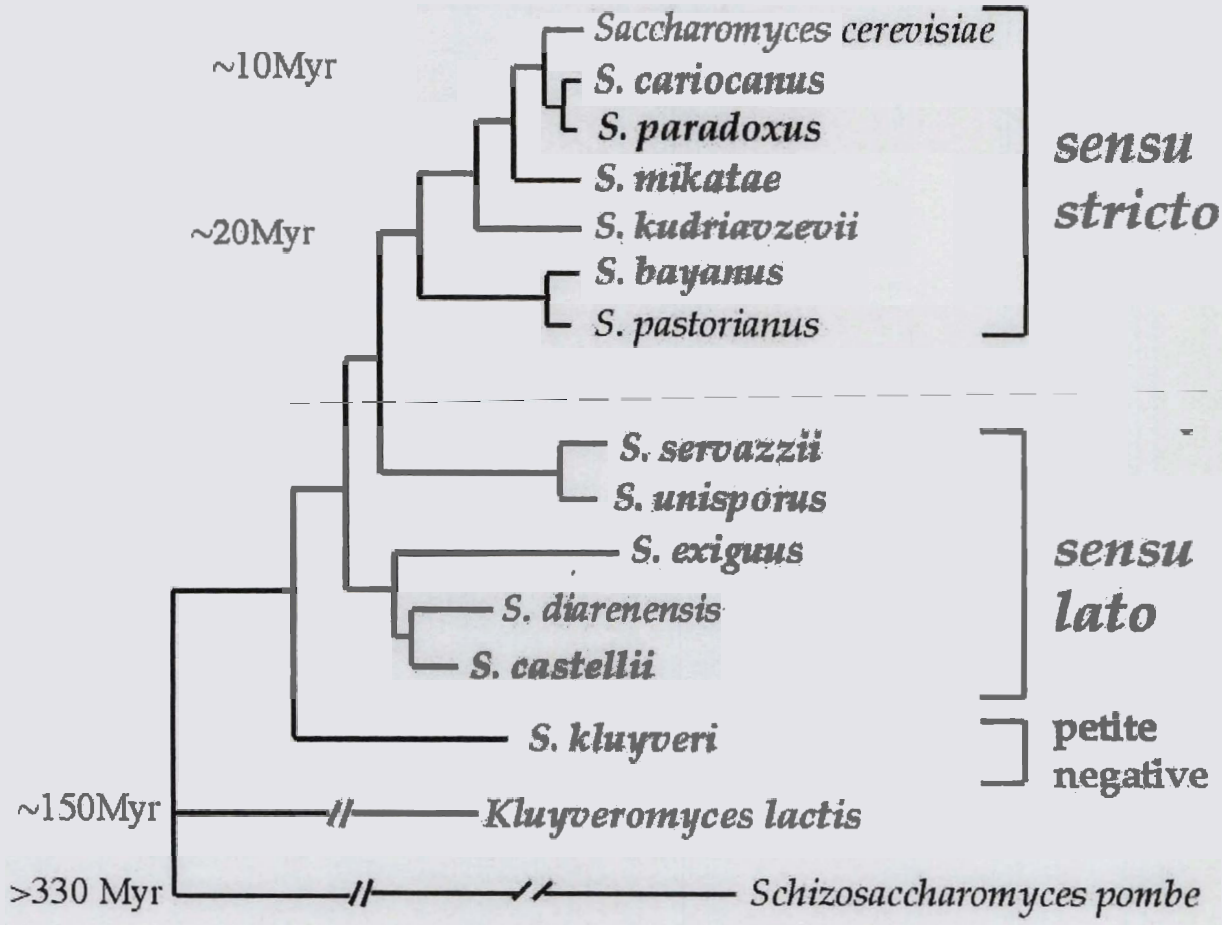

Figure 5: Phylogeny showing the relationship between *S. cerevisiae*, *S. kluyveri*, and *S. pombe*. From Cliften et al. (2003).

Figure 4

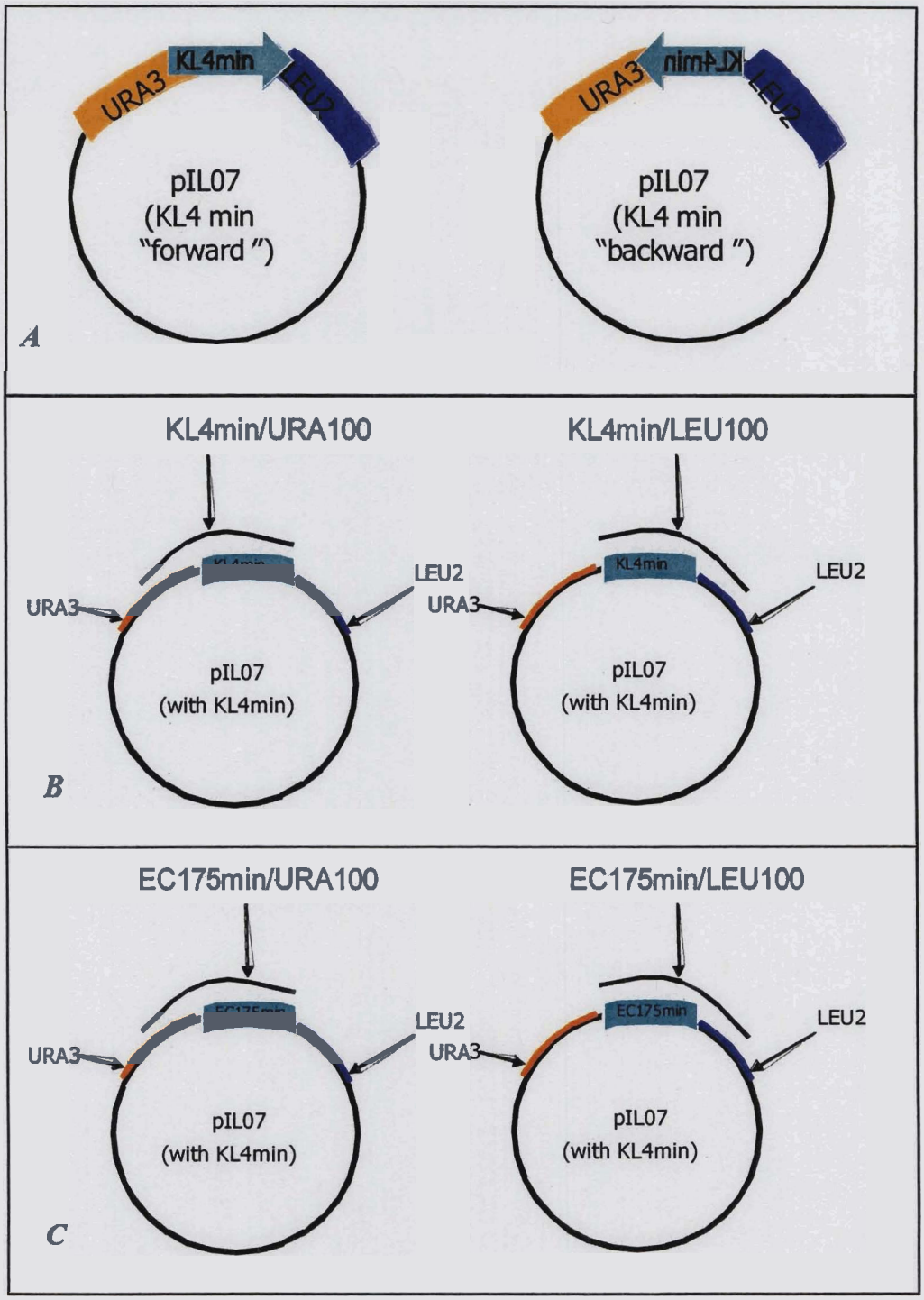

Figure 4: Diagrams of pIL07 with inserts (genes not to scale, but used for reference to determine "forward" and "backward" orientation of ARSs). **(A)** Diagram of "forward" and "backward" orientations of a truncated ARS. **(B)** Diagram of amplified PCR products from pIL07 containing KL4min. Amplified sequences were inserted into pRS406 to test for ARS function. **(C)** Diagram of amplified PCR products from pIL07 containing EC175min. Amplified sequences were inserted into pRS406 to test for ARS function.

Figure 5

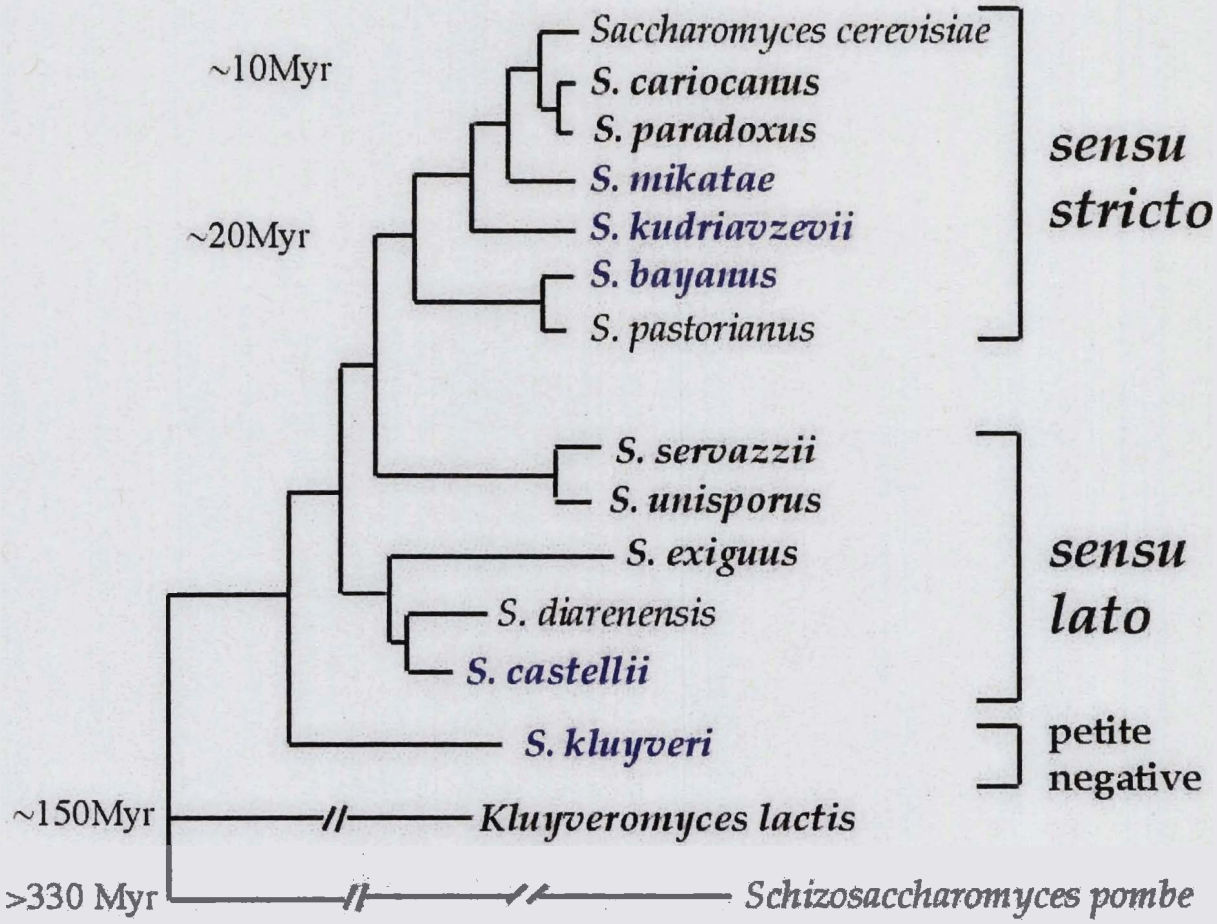

Figure 5: Phylogeny showing the relationship between *S. cerevisiae*, *S. kluyveri*, and *S. pombe*. From Cliften et al. (2003).

Table 1

Yeast Species and Strain	Relevant Genotype
S. cerevisiae W303	*leu2-, ura3-*
S. kluyveri FM628	*ura-*

Table 1: Yeast strains used in this study. *S. kluyveri* FM628 was obtained from Mark Johnston's lab at Washington University.

Table 2

Plasmid	Description
pIL07	Contains URA3, LEU2, a BamHI site, lacks an ARS
pRS406	Contains URA3, a BamHI site, lacks an ARS
pIL17	Contains URA3, a MscI site, contains EC175min (functional ARS only in the presence of a flanking element)

Table 2: Backbone plasmids used in this study as vectors for inserts. pIL07 was used in the ARS screen. pRS406 was used to determine the location and function of the flanking element discovered in pIL07. pIL17 was used in the ARS and flanking element screen.

Table 3

	Primer	Purpose	Sequence
Sequencing	IL325	Sequence inserts in pIL07 (used with IL326)	GCCAAACAACCAATTACTTG TTGAGA
	IL326	Sequence inserts in pIL07 (used with IL325)	TTCGTTGCTTGTCTTCCCTAG TTTC
	IL429	Sequence inserts in pIL17 (used with IL430)	GTGAGCGCGCGTAATACGAC TC
	IL430	Sequence inserts in pIL17 (used with IL429)	CACACAGGAAACAGCTATG ACCATG
Truncations	HP171F	Amplify HP171 – full length left side	GATCGAagatctCATTTTCGAA ATTATTAAAAATTAAGG
	HP171R	Amplify HP171 – full length right side	GATCGAagatctGCCTGTCAAG CTCTTCCGAATG
	HP171-250L	Amplify HP171 – truncated left side	GATCGAagatcTAAAATCTGTA TGTTTACCTATTTAAATCAC
	HP171-250R	Amplify HP171 – truncated right side	GATCGAagatcTGGGCAAACG TAAGAATTTGATG
	AC16F	Amplify AC16 – full length left side	GATCGAagatctAGTAAATATA GAGTAACAAAAGACGCTG
	AC16R	Amplify AC16 – full length right side	GATCGAagatcTCTGTTAACCT GCAAAGTATACCTT
	AC16-300L	Amplify AC16 – truncated left side	GATCGAagatcTAATCACGTGA CCATATAGTTGCTAGG
	AC16-300R	Amplify AC16 – truncated right side	GATCGAagatcTAACAGCTCAT TGGTTGATGGAGAC
	KL4F	Amplify KL4 – full length left side	GATCGAagatcTTTAATACTAT TGACCAATTGTTTAATTA
	KL4R	Amplify KL4 – full length right side	GATCGAagatctAAAGTTGATG GTTTTTTGATTAATTAAAG
	KL4-100L	Amplify KL4 – truncated left side	GATCGAagatcTTCATAATTTA CCAAATGGAATTTT
	KL4-100R	Amplify KL4 – truncated right side	GATCGAagatcTTTTCAACAAT AATAGTAAAACCCGC
	KL4-92L	Amplify KL4min (used with KL4-92R)	GATCGAagatcTTCTCCATTTC ATAATTTACCAAATG
	KL4-92R	Amplify KL4 min (used with KL4 -92L)	GATCGAagatcTACTTTTTTTTT TCAACAATAATAGTAAAAC
Mutagenesis	KL4 mutL1	Introduce mutation L1	GATCGAagatcTTCTCCAgggCA TAATTTACCAAATGGAATTT TTTT
	KL4 mutL2	Introduce mutation L2	GATCGAagatcTTCTCCATTTC ATAATTTACCgggTGGAATTT TTTTTATGTTTCCTATAAT
	KL4 mutL3	Introduce mutation L3	GATCGAagatcTTCTCCATTTC ATAATTTACCAAATGGAATg ggTTTTATGTTTCCTATAATC TACGCGG
	KL4 mutL4	Introduce mutation L4	GATCGAagatcTTCTCCATTTC ATAATTTACCAAATGGAATT TTTTTTATGgggCCTATAATCT ACGCGGGTTTTAC

	KL4 mutR1	Introduce mutation R1	GATCGAagatcTACTTTTTcccT TCAACAATAATAGTAAAACC CGC
	KL4 mutR2	Introduce mutation R2	GATCGAagatcTACTTTTTTTTT TCAACAATAcccGTAAAACCC GCGTAGATTATAGG
	KL4 mutR3	Introduce mutation R3	GATCGAagatcTACTTTTTTTTT TCAACAATAATAGTAAAACC CGCcccGATTATAGGAAACAT AAAAAAAATTCC
Isolating the flanking element from pIL07	EC175minL	Amplify EC175min (used with EC175minR)	GATCGAagatctGATACGCATC ACTATGCAATTGTTAG
	EC175minR	Amplify EC175min (used with EC175minL)	GATCGAagatctGACCTTCCAG CGTAAAAACATTAAAC
	IL431	Amplify truncated ARS plus 100 bp of pIL07, *URA 3* side (used with KL4-92L and EC175minL)	GATCGAagatctAATAATGCAA ATAAACATTTGAAaTTATTG
	IL432	Amplify truncated ARS plus 100 bp of pIL07, *LEU 2* side (used with KL4-92R and EC175minR)	GATCGAagatctACCAATTACT TGTTGAGAAATAGAGTATAA
	IL433	Amplify KL4min, add MscI restriction site on right side (used with KL4-92L)	GATCGAagatctggccaTACTTTT TTTTTTCAACAATAATAGTA AAAC
	IL434	Amplify EC175min, add MscI restriction site on right side (used with EC175minL)	GATCGAagatctggccaGACCTTC CAGCGTAAAAACATTAAAC

Table 4

Insert	Description	Maintenance
KL4	Full-length KL4	+
KL4min	Truncated KL4	-
EC175	Full-length EC175	-
EC175min	Truncated EC175	-
KL4min/URA100	Truncated KL4 plus 100 bp from pIL07, *URA 3* side	+
KL4min/LEU100	Truncated KL4 plus 100 bp from pIL07, *LEU 2* side	-
EC175min/URA100	Truncated EC175 plus 100 bp from pIL07, *URA 3* side	+
EC175min/LEU100	Truncated EC175 plus 100 bp from pIL07, *LEU 2* side	-

Table 4: Results for isolating the flanking element from pIL07. Inserts were cloned into pRS496 and tested for ARS function. See Figure 4 for diagram of truncated ARSs plus sequences from pIL07.

Table 5

Insert	Sequence
pIL17-AluI-2	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGTAACTACAGTTGATCGGAC GGGAAACGGTGCTTTCTGGTAGATATGGCCGCAACCGACAGAAATAGGAAAAAAAC ACAAGGACATGGAACAAAACAAAACATGTGACACAGAACAAGGAAGGGTTATTTGT TTACATTTTCATAGGATATTTGTTTACATTTGGAATATATTTCACGTGACCAAAGA GGATAATAAGTAGAAAAGTAAAATTAAAAGGGCAGCGGAAGGGCTCAAACAATTCG GAGGAGCACGGAAATATAAGCAGAGATAGGTTACTAAATTTTAGTGGTTAGTATGT GTGAAATGGAGAATATTTCGAAGAGCAAGATTTGAATTATTGAATAGTTGAAAATA AGTAACATTCATTAATGTTTAAAGAAAAGTATTAATAACATAAAGGTTAAAAATGA GAAGCCA
pIL17-AluI-3	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGTTCGCCTTCTGCCCTTCGT AAAATCACTAGGTTTTCCCTCTCTACGAAACATATACATTTGCAATGGCTTTTCAC AATCCTCTTTGTTAATCTTGTAAACCAATTCACCTTTGACTCTGTTAAACCTCAAA GAATGATCACCAAGTTTTAGAAGGAAATAAGTTCCAGGATGTAAGCCAGTTTTTGA AAAAAAGTCTTTGATCAAAATCACCAAAGTTCCATCTTGTATAATCATAAGTAATC GCCAGTTTATTGCTGCTTCTTTGCTGCATCCCAACTACTGCTAACTTTTATTTATT TATTTATATATATACATTCGGCTTACTGTATGTATAAGCCATCATTCGCTGCGCCG TCGATAGAAAATTGAGATTAACCGGAAAACTAATAATTATCACAAGAATTTACGTA GCGAAAAATAAGAAAGACAATAACCAGAATTTCTCAGGATAATCATGGAATGCAAG AGTAAAACAATCATCCATGTATCGTAGGCATGCTTCTGGAGCCA
pIL17-AluI-4	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTATTCCCAGCGTCCAAGA GCCCGAGAACTTCAGGCCTGAAGGCAATTCCATTGCCGTTGGTGAGGATGATGACG TGATTTAAAGTTTCTGTTCGCACTGACAGTAGCGTTCCAAATATACTTATATGTAT AAAGGCTTTATTTTGAGCCATGTATCTTTTGTACTACACTCAAACTAAACAGCATT CCGGTGCATTAAAAGCATCTTTCTTTACATTCAACTACAAGAAAGGAAAAAAAAAA AAAGATAAAAGAGGAGAAAGAAAACAGGAAGAGAAACATTATCCTGTCTTAACCTA ACCCGATGATTAATTCGAAACTGGGTATGAATTATACTTCTTTATTAATGTAGATT TAGCCTTTTCCTTTTTTTAATTTTTTAACCTTCCATTGTACATTAAAATTAATTTT TAAAATGCTAGACCTCATAGTAAATGACAACAACTTAATACTGTAGATATCCTACA TTTCTCAAGAGAGATGCTCTGTGAAGGAAGGAGAACAAGGGCGAAAGATAGTGCAA CAGTTTTTTACTCTGACGCACACTTATACTGGTATCCGTCTCGGTAATAGATTGTG TGTATCACGTGACTCTTTCGTTTCGAATGAGTATCACAGATTAAGCCA
pIL17-AluI-5	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTTTTTTAACCAGAGTGTT GGGGTTTGGAACTGTGCGCATGGACTACTTTCAAATTTCTTATATTGATTACCAAC ACGCACCCAGTAATAAGATCCCCTGTATGTATTTCGTTGGTTTATATTTGTTTATG GTCACGTGTGCTCATCTACAGTTTTCGTATTTTCTTGGAAACTCGATTATGAGGAG TTGAAATTTTCAAAGTGGTTCGAAATTGCACGTACAGCAGCGATATACTTGGACAG TTTGCCTTGCTTGGTTTCTTATTTTTTTTTTCTTTCTTGTTTACGCTCAATTCCTA CAACGCCTGTTCCCAGGATACGTTGAATTACTGTCACCTGTCTAGTCTAAAGCCAT AAGAAATAAACAAAGCATGTAATCGAAAAAAAAAAAAACAAAAAAAAAAAACAGGCAA CCATGAAATGGATCCTAGTCTACTGCAACTTTCAAAGCAAGGATTTGCAAGACATA ATAATGTTTATGACCA
pIL17-AluI-6	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCAAAGTAATATTAAGAAAC AATAAATTTTGGTCAAATACAAAATCTTGCTTTAAGCCCGAACAAATCTCCAAGCA GGGACACCTTGGTGATACACCCCGATCGTTATTAGAAGTGAAGTATTTTCAAATTG CATACAATCCATTCCAGAAATAAGTTAGGTTCTGTTCATACAAAAAGAATCAAGCG GCATACTTTGTCTTTGGCTATCAACTCTTCACTAGTTTTTATTTTAGTTAAAAAAT AACACTAATAATATTCCCAAGTAAACAAGACTAGGATTATTCCA
pIL17-AluI-7	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTAAATATTTTTCTGCGAA

	GAACAAAATTAAAGAAAAATTGAGAGAATACACTGCACAGTGTGATAAAGAAAAAG AAAGATGAAAAGAAGTAGTGATAAACTTTGGGTATCTATTGCAAGTAGACGGTCCG GGCAATTGTTACTTCGCAGCCCATTATACACGCCTGCTTTTAGGAGAAAGTCATTC ATGTTAAATTGAAGTAAAGTTAGTAAAGAATTAATGATTAGCAATAAACGATGAAA TATATATACATATATATGTGTGTATGTATATACACATGTGTGTATGAATAGCCGTG CCCCATTTCTTCTGATATTATACTAGCGAATGTTATTTTGATACGTTAGATGCTTT AGCGTCCAGTCAGCAAAGTTCCTCTTAAAGCAATTGCTAAAAGCCAGGATCTCAAG AAAGATTCAATAGAGTGGGGCGAAAAATAGTGATAATAGCAATGTTCGTTCTGTAT AACGAATCGAACAAAGAAATAAAAACTCCCGTTGATGATAGTACAATCAGAACCGA TTTCTGATCCGACGATATAGTTGTCTAGTCTCTTGTCCCTTTTTTTTTTTTTTTT TTTTAATATACTGTGTTTTCTGATCGAGCCA
pIL17-AluI-8	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTTCAATTAAAAAATTGGA GCCTCTTCCAACAGAAGAGTTAAAACCATTGCACGATATTAAACCCAAAACTAGCG GCGGTAGTAGCGTTAATTTATCACTAAACCCAACTACAGTTACGAAAAGTGGAAAG AAAAAAGTTGTCCCGACCACTATACAAAGTACCTCGATGGAGTTTAACCCTCCATC TTACAGTGTGCCTAAAGATCTCAAGAGAAAACCAAAGGATGAAGAAACGATGGTAA ACGGCAAAAAGCAAAAAAAAGATGTAGATCCAATGGACTTTCTCGATGCAAACATT CTAATGCCTAATATTTCTTTTTCAAAGTTGAGGTTGGCTACCCCAAAAATTCGTGT GTCCTTTCAATACTCGCCAGTGCATAATAAGAATTTAATATTCGACGTCAAGAATG GCACAGGTAATGAACAGAAACCTACATCTATATCCTTACTTTTAAAGGAGCCAGAG CAAGATAGAATGATATTTCAAGATTTTATACCAAAATTTGTCACATTATCTAGCGC CGGAGAGAGTTTCTGGGCATGCTGCACAGAAGATGGCACACTGTACGTTTACTCAG ATACAGGTAAAAGATTGCTTCCACCAATGATTATGGGCGTGCCATGCAGCCA
pIL17-AluI-9	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTGTGTTTAAATGTGATTG CTAGGATTATATTTAAAATAATTTTATTATAATATAAATAATAATAATATAAATAA TATTATATATTATTAAATAAATATCCAATATTTTTATATTTATTAATTAGAAAAAT AAGTTATATTTAAATATAAATAATAATATATAATATATAAATATAATATAATAAAT AAATATAGAATTATATTATTTAAAAAATAATTAATTAATTAATTAATTAATAATAA TTAAATATTCTTGATAAAGTAAAGCAGCACAATTAAACACAGCCA
pIL17-AluI-10	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGAATCTTGGGTAAACACGCG AAACTAAAGAAAAACAAAAAAAATGAAAAAATAAAAGATACAAATTTGAAATTAGC CA
pIL17-AluI-11	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGTAGCAAAATAAGATAATAA TTCAATTAAAACTAATAAAGGAACTAAAGGTAAAGGAGTACCTGCAGGAACAAATA ATGAGAAGAAAACTCAACCATGTCTATATAAACCTAAAATTGTTACACCTAATCAA ATAACAGTACTTAATGATACAACAAATACTCCA
pIL17-AluI-12	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTGCGACAAAAAAATATGA AAAACCGTTACTGGGTGATCAAAAAGGAGCCA
pIL17-AluI-13	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTACGTACAAGGTTCACGG TGTTGCAGACAGAAATATTCTAGAATTATACTTCAATGATGGTGTTGCTGTATCTA CTAATACTTTCTTCTTCACTGAGGGCAACTTCATTGGTGAAGTTGACATCACCACT GGCGTTGATAATGTGTACGAAATTTTGGATTTTGAAGTAAAACAATTGAGTTCAAA ATATTAAACGCATAAAAATATTTATTCAAATTCCAAATTAGGTGTCAGAAATATTT GAATACGGAATGATCATCTTTCTCTACTTGTTTTTGTTATGTACGAAATAGTCTTT TCATGATTAGTATGTACTTTGATAAAAAACTTTTTTTTACTGAAAAATTTTAAGTGA AACTTGCATTTGTTGTATTATTGATTTAAAAAACTGTCTATGATACTGCTTCTCAT AAATGGATACACTATCTAGCCA
pIL17-AluI-K1	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGAACGATACCTGGCAAAAGA AATGTATTTCCACAAAACTGAGTATCAGGTTTAATTATCAACTAGACTTTGCTGGG

	TATGTGGATATTTCTTCTACTCTCACTCTTCTGGAAATGCTTTTTTTTTTTTTTTG ATGTTACATATGTATTTAGAAAGAATAGTTGATTCATGTTCAAAAACATGAAATTA ATTCGAGCCTTCTAATTATCAGAGTTCTTTCCCACAAAATAAAACAGCAGCCTTTA AAATATGACAAGTCTAGGTCTGGAAGGAATATTAATACGTTTTCTTGGAGCCTCAA GGCAAATGCAATGTTGATCTGAGAAAACCTAGATGTAATAATGAATAAATATTTAA GAAAAGTATGTAAAAAAAAAAACATCGATTCCTAGAAAATTGGTATCGGTGATGCA AAATACTATGAATCAAGAGTTTTTCAGAGCCA
pIL17-AluI-K2	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGAAAGGATCTTTTTATATTA TTAAAAAATCATCTAATCCTTTAAGATATATACATTCTTTTATATTAATACTAAAA TTAGATTATATTGTTATATATAGTATTAAATTATTATTAAATATTAATTCTTCTAT TATTATAAAATCATTTAATAATATAAATATATTTTATATATTAAAAACAACTAATA ATTCTAATATTGAATATATTATTAATTATTTTATATATAATAATTATAAATCTATA TTTTTAGGTTTAAAAAGTTATGAATATAAAATTTGATCTAGAACTTATATAAAATA TAAAATAAATTATCTTAAATTATTAAAAATTAATAATTTATTACTTAAATATAAAT AAATAATTATATAATATTTATATATTTGGTATAATAATATATTTAAACATAATTTT TATAAAATAAATCATCTAAAAAATATTATATTTCTTTTTATATTAATTAGATAATA TATTATATTAATTAGATAATATAATATTTTAATTTAAGGAATAGTCCGAACAATAT AGTAATATATTGAATTAACAAATAAGATATCCTCCTTTATCATCTATTCAAGCACA TTCAGGACCTAGTGTAGATTTAGCAATTTTTGCATTACATTTAACATCAATTTCAT CATTATTAGGTGCTATTAATTTCATTGTAACAaCATTAAATATGAGAACAAATGGg TATGACAATGCATAAATTACCATTATTTGtATcC
pIL17-AluI-K3	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGTAGGTAGTTTATATAGTCC TCTAATGTGGAAAAACGCTTGTAAATATTTAATTTTGTGGAGGTCCCTCTCAATCA TTTGAAAAAACTTCGAACTAACTTTTCTGTCTTACAACAAGCACCCCGTACAAAAG ATAGAAAAGGCAGCAATGATTAAACCCAAATTCAGCATAATTCGTTCTCTAAGTTA CTCTGCTCGTCGTCTGAGACCACAGCAACGCATCGAAATCAAAACATTAGAGGATT TGGCTAAATTAAAGTCTATAGATGATGTCGACCCAGAGTTGATAAGGCAACTAATA AATGAGCGTACCAATGAGTTGAACATAAAAAACGAGCCA
pIL17-AluI-K5	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGTAACTACAGTTGATCGGAC GGGAAACGGTGCTTTCTGGTAGATATGGCCGCAACCGACAGAAATAGGAAAAAAAC ACAAGGACATGGAACAAAACAAAACATGTGACACAGAACAAGGAAGGGTTATTTGT TTACATTTTCATAGGATATTTGTTTACATTTGGAATATATTTCACGTGACCAAAGA GGATAATAAGTAGAAAAGTAAAATTAAAAGGGCAGCGGAAGGGCTCAAACAATTCG GAGGAGCACGGAAATATAAGCAGAGATAGGTTACTAAATTTTAGTGGTTAGTATGT GTGAAATGGAGAATATTTCGAAGAGCAAGATTTGAATTATTGAATAGTTGAAAATA AGTAACATTCATTAATGTTTAAAGAAAAGTATTAATAACATAAAGGTTAAAAATGA GAAGCCA
pIL17-AluI-K9	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTACTTTCAGGACACAGTT CCGCTGAAGAGGAATCTATAGCCCAGTTAAGAAAACGTTTATTGGGTAAAAGGAAA GGAAGTGTGTCCCTAGATGCAGACGGCTCTTCTATCGACAAGCAAATGCAAGTTCA AGATGACCTACAACAAAATTTGATCCAAGATATGTCACTGCTCCTCAAGAAACCAG ATGGGACAAGTCAACTAGGGCTTTGGCTCATAACTTGGCATCTCAAATCAAGGGCA ACGAATTACCAAACCAAGAGTTGATCGGACGGGAAACGGTGCTTTCTGGTAGATAT GGCCGCAACCGACAGAAATAGGAAAAAAACACAAGGACATGGAACAAAACAAAACA TGTGACACAGAACAAGGAAGGGTTATTTGTTTACATTTTCATAGGATATTTGTTTA CATTTGGAATATATTTCACGTGACCAAAGAGGATAATAAGTAGAAAAGTAAAATTA AAAGGGCAGCGGAAGGGCTCAAACAATTCGGAGGAGCACGGAAATATAAGCAGAGA TAGGTTACTAAATTTTAGTGGTTAGTATGTGTGAAATGGAGAATATTTCGAAGAGC AAGATTTGAATTATTGAATAGTTGAAAATAAGTAACATTCATTAATGTTTAAAGAA AAGTATTAATAACATAAAGGTTAAAAATGAGCCA
pIL17-AluI-K10	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGTATTTTTTTTTAGATATAAA

	TAATATTATCATTAATTTGATATTTTTTTTTATAAATAAATTAATTAATCCTCTTTG TTTATTAATATAAAGAACCAAAGAAATTATATTTATTTTATTATTAAATAAATTCA TTATTATATATTTTATAATCATAATTAGAAATATTATTAAAATTATTATTATTATT ATATTAGATATTTTAGATTTATATAATATATGTATATTATATTATATATTATATAT TATATATTATATGTTATATGTTATATTAGGGATTCGAACCCCAAACCTTTCGATTA CAAAACGAACGCTCTACCAATTGAGCTCCAATTTGTAAGCGACGGACCATGGCAGA ATAATTGAATCGGTGTCTTCTAAGCCA
pIL17-AluI-K13	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTTCTCATTTTTAACCTTT ATGTTATTAATACTTTTCTTTAAACATTAATGAATGTTACTTATTTTCAACTATTC AATAATTCAAATCTTGCTCTTCGAAATATTCTCCATTTCACACATACTAACCACTA AAATTTAGTAACCTATCTCTGCTTATATTTCCGTGCTCCTCCGAATTGTTTGAGCC CTTCCGCTGCCCTTTTAATTTTACTTTTCTACTTATTATCCTCTTTGGTCACGTGA AATATATTCCAAATGTAAACAAATATCCTATGAAAATGTAAACAAATAACCCTTCC TTGTTCTGTGTCACATGTTTTGTTTTGTTCCATGTCCTTGTGTTTTTTTCCTATTT CTGTCGGTTGCGGCCATATCTACCAGAAAGCACCGTTTCCCGTCCGATCAACTGTA GCCA
pIL17-AluI-K16	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTTCCCATCATGTGTCCCC TGCTGATTTGTTTGTGTAATGTTATTTTTTTTTTTTTTGGTGCAATTCTTTGACTTT CAGTCTGTAAGATAATAAGTTGTGGGATAACAAAAACATTTATAAACGACAAAAAA CTAAAAATTCTTGAAGACGGAAAACAGAGGAATCACTCAGCGGTTGTGTATTATTC AGAACTTCTTATTTGGTAGTTACATGATAATATAAACTTACCAACACTTTAATTGT TAAAAATCATCCCTTCAAATATCCATGTTGTACCTATATTCCTTTCACGTGACTGG ATTCTAGTCTTCAGCATATCATGCCTGTAATATGCGACATGGGCAAAAAGTTGAGA AATCTGTTGATATTGGGACAATTGTTGAGATTCCATTTATGGTTAAGGCTATAATA CTTTGGTATATGGTTATACTAGAAGTTCTCCTCGCCTTTCTAGGAATCCAGAAAAG GGAGTCGATAATTCTACATACAAATATCCTTCCTTTTATATGTTTCTGGTTCATTA TCCTATTACATTATTGTTGGAACGAAATACAACTATCGACCATCGACTAGTATTCG TGTTACTAGTATATTATCATACACGGTGTTATAAGGTGACATAAAGAATGAGAAAC AGTCATCTAAGTTAGTGGAACCA
pIL17-AluI-K17	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTTTATTTTTCAATTATAA TTATGCTGCTTGTTTATAGGGTAGAATAGTTAGACAATTTTTCTTAGGTTCAAAAT AAGTTCTAAACTACGTTTTTACTAAAAAATATAAACAAAATTTTAAAATTGAGGCG GCAGAAAGAGATTAGAGTAAAATGCGCCTTCCGATCCTATGGGCATTTGGTAGGTA ATCGGATAAGTAGGTATATGAATATCCATATTGTTTAAGACTAATAATCATGTTGT TGGAATGTTAATCAACTATGATATATTGACTAGTGCTCGTATTGCTAGTATAATAT CACACACGGCGTAAGGCGGATGACATGAAGAATGAAAAATAATCATCAGATTTAGT GGAAGCCA
pIL17-AluI-K18	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGTTTGAAATTTATGCCAAAC CATGAATAGAAAAAAACTGGAGAGTATACAAAAAAAGCACGACTGGGGTTTGAAGC AGTTTCCAAATTATGTAGGGTCTGAAATCGAGCCTTTTTTTGTCCTATTATGGATG ATCCA
pIL17-AluI-K19	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTTGAACAAGAAAAGAAGG AATCGAACAAGTGAATGCTGTTGCCAACTCATTTTCTTTTCAAGATTTTGTATATT AACTTTTTTTTGGCATGTAATATAAAGATTTTCTTAGAGGAAATATGAAATCGAGA TATATAAAGTCATTCGCATTTATTAAGGAACTACTTGTTTCATTTAATTTTTATTC TTTTTAAATGATATTTTCGTCTCCTAAAGTTGATGAAAATTTCGAAAACTGGAAAT TACTAGAAGAACATGAAGAATAAATTCAATCATACTAGTATTATAGTTATATAATT AATCCATACAAACGAATAAATTACCTATCTCTGATACGGCCACAGCCAATCTAATA CCAAATTCTGCAACAAAGAATAGAGTTGCGAACTAATCCAATAAAGCGACAACAGG ACAGAAGCCTGAGAGGAAACACCCATTAGAAAGATGCCA
pIL17-AluI-K20	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG

	CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTCGTATCAGTTTTATGAG GTAAAGCGAATGATTAGAGGTACTGGGGTTGAAATGaCCTTAACCTATTCTCAAAC TTTAAATATGTAAGAAGTCCTTGTTACTTAATTGAACGTGGACATTTGAATGAAGA GCCAGATCTTAAAGTAGACATACTGCGTCCCGCTATCGGGAAAaCCTTATTTGATG TCTCAATTTGTGCAAATCTTTTCCAATGAGGGTCAAGCATACTATCATGGCATAGG CTTATCGGTAACGTaCATGAAAAGTCAGAAGGATATCAAAAAAAAAAAAAAAAAAA TAAAAAAATAAAAAAATAAAAAGCAAAAGAGAAAACATTTAGTGCATTAATATTAT TTCATAACAGGTAGGAGGCGGCGGGTGAGGTGAAAGAAAGAGAAAAAGAAATACTC CTCGGGGACGTAACTATACAAAAGTGGATATTGAGAAGTGATTCTTTAAAAATTTC GTTAGGATTATACAGTTGGTGCGTTTGGCATGAAGTCCA
pIL17-AluI-K21	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGAATATATGTAAGTATAGCC TTTATAAATATAATTATTAATAAAAATAATATAAATTTATTAAAGAACGGTCCA
pIL17-AluI-K22	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTTCTATTGAAAATGTTAT TTTCTGAAGGAAAAATATCCGTCTGCAACAACAGTTCTCTACCAATGCTTTTAAAT ATGATCTTATTACAGGAAGAGTTTGAATTGGAGTTTAAAATCAAGTTATTTCAGAA CTTTTACACGATACATAGGCACTTAAACTACGAGGATACGGTGACGATACTATTCA AAAATTGCAAAGGAAATGGATATCGGTCTATAGAGCCA
pIL17-AluI-K24	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTTACGACCCAAACACAAA ATTAATCAACGCCGCTGATGAAGAATCAGACGACTTTTTGACAAATGTTGAACAAA TCAACAAGTTGACAAAGGACTTAATCGCTGAGTCGAACCCAAACTTTACTCCAAAG CCAAACGACCAGTCCACTCTGATGATTAAAAAGCCA
pIL17-RsaI-1	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGATGCCCATCATTTCCACTT CTACCTACGATAACCAGCAAAACTCTTATAACGTGACCAAAATTTTAAATCCAGAT TTTTCCATCAATTTGGAAAAATACAAGGGCTATTCTCCACTTTACGTTCCATTCTC ATATTTATTGTCGTATGCTTTGAATTTTGCAGCAGTCACTGCAGTCTTCGTCCATT GTGCACTATACCATGGCAAAGATATTATTTCCAAGTTGAGAAACAGCAATCATGGC GGTTCAGACATCCATATGAGGCTGTATTCCAAAAACTACAAAGACTGTCCAGACTG GTGGTCCA
pIL17-RsaI-2	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGTAGCTAACGATACCTGGCA AAAGAAATGTATTTCCACAAAACTGAGTATCAGGTTTAATTATCAACTAGACTTTG CTGGGTATGTGGATATTTCTTCTACTCTCACTCTTCTGGAAATGCTTTTTTTTTTT TTTTGATGTTACATATGTATTtAGAAAGAATAGTTGATTCATGTTCAAAAACATGA AATTAATTCGAGCCTTCTAATTATCAGAGTTCTTTCCCACAAAATAAAACAGCAGC CTTTAAAATATGACAAGTCTAGGTCTGGAAGGAATATTAATACGTTTTCTTGGAGC CTCAAGGCAAATGCAATGTTGATCTGAGAAAACCTAGATGTAATAATGAATAAATA
	TTTAAGAAAAGTATGTAAAAAAAAAAACATCGATTCCTAGAAAATTGGTATCGGTG ATGCAAAATACTATGAATCAAGAGTTTTTCAGAGCTATAAGTGAGAGAAAAAATCC CTAGAATTTAATTAACAATAAGCACACGTAATTTATGGGCATGTTTATTAAATGCA TGTCACTTTCGCCAGATTATGTGCTCACGCTATCAGGAAATTCACACGTGCCTTTA TCTGTATTTATCACAGCCACTTCACCGTCAGTATTAATCTCGAAAGATGCTTGTTT TTCTGGTTGTTCTTCCTGTAGGTATTCTTCAGAAGTAAACGTTTTCAGGTTCTCAT CAATTTTGTAAAGTTTAACACCCTCACTGATCTGCTTTTCTTTACAGATAGAGCTG ATGTAAGACGCCAATCGCTTAATATCGTTTTCGACATTTTCACTAATTGACATGTG GTCTTCTGGGTATATTAAAAAGTCACCGTTCTCTAAAGTTTGGTAGCTCGGATTTT GTATCAAAGTGAACTGTAAGTATTCCGCCAAACAATCACTTAACTCATCTTTTGTA ACACCGGGTGCAACGCCTTTGACTAGAAAACAGTTAAATTCAGCGTTATGCGGCCA
pIL17-RsaI-3	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGACGAAATTTTGGATTTTGA AGTAAAACAATTGAGTTCAAAATATTAAACGCATAAAAATATTTATTCAAATTCCA AATTAGGTGTCAGAAATATTTGAATACGGAATGATCATCTTTCTCTACTTGTTTTT GTTACCA

pIL17-RsaI-4	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTACTACAGAAGTATCAGT ATCTAATAATCTTAATTGACCATCTTCTAATAAATCTTCAGGAATAACATATGATT CAAATTCTACTGTTTCACCACTTTCATTAATAAAATCTGAATATTCATATTTTCAA TATCATTGTAAACCAATAGCTTTAATAGTCATAGCAGGTGAAATAACTTCATCACA TAAATATAATAAAATAAATGAAGGAAAAGCAATAATTAATAATACTACAGCTGGAA AAATTGTTCAAATAATCTCAATTGTTTGACCATGTTTAATATATTTATATGCAATT GGATTTTTTGAATATGTTCTTAGAATAGTAAATAATAATCATGATACTAAACCTAA AATTACTAATAAATAAAACATAATATTATCATGTAATTCTAAAATACCTTCTTGAT TTGGTGTTGCTGAATCTTGAAAATAAACTCCATAAGGTGTTGGCCA
pIL17-RsaI-5	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGGTTCTTTAATAAATTTATA TTATTTTTATTAATAATTATATTTATAAAGGCTATACTTACATATATTTCATAGAA AACCAGCTATCTGCAAATCAGATTTGTCTTTCACCACTTATCACAACTAATCAGAT GATTTTGCAACATCAACCTGTTCGACCGTTTTATTAACTTTTTTATTTATTAATTT TTTCCGTCTTGTCATAATAAGATCATTTGCTTTCGGGTCAATTAATATTAACTTTT TACACAATTTTTTTTTTATTCAATTTTTTTTTAATTTTTTTATTTTGCTAATATTA ATTACTTGCTGACCCATTATACAAAAGGCCA
pIL17-RsaI-6	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGACAACTTTTATAAAAATAA GATTATCACCACGTAAGGCAGGACGAAACACCAAAACATAGAAACACATACCTGCA AAGGTATGAACGATATCTTCTATAAGCCA
pIL17-RsaI-7	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCGACGCAATTCTGGCCCGT GCGAGGGATTCTGGTGGAAATGATTGGGTAATTTTTAAACATATATTTTTTGAGCG TGCGCCACTAAAGAGATCTATTAAGAAGCATGGAATAAAGAAGGGTAGTGAACATC GCATATTTTCGGAGTCCA
pIL17-RsaI-10	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTGAAAAATCTCATTATTA GATCACACTAAGACAGTATCTGCAGTCTATTACTTTAATGATACTTTATTGTAGAA ACGAATAATAAAAATCTAGTTATTAGAAACAATCCTGTTTAACGCGTCACTTTTTG CTGTTATAAGAGGCCATCTTCAAGATATCCAATCTATTGGTTCTTTTATTTTTTTT AAATAACTAGCTAAATTTATATTTTTTATTGCGTGGTAAATTTGGCTCTAATAGGA AAAAAAGTGCCCTAAAAATCAAGTAAATAAACATATTCTTCAGAAGCATAGAAAAA CAACTTGACACCAATTGAAAGGGATGCGAGCCA
pIL17-RsaI-11	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCCACAAGTATTAAAAATAT ACGGACGAACTAAAAATATAAAATTGTAGATTATGAACAACCATACGGGCATGCAG GCGAGAAAGTAAAGACAAGACTTTTCATTTCGTTGCCA
pIL17-RsaI-12	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCCATCATGCCTGCATTTTT TATGAATTCAAATTTAAATCATTGTGTTTTAATTGAAAGTTAAAAGCTATATTTGA ATAAAAATGTATAACATTAAAAAATAAAAAACATAATAACAAAGATAAAAAATCAT TTAAAACAGGTTAGTTAACTTCAAATTGAATACGCTCTATTCAAAAAGTGTCCA
pIL17-RsaI-13	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGACCGTTCTTTAATAAATTT ATATTATTTTTATTAATAATTATATTTATAAAGGCTATACTTACATATATTTCATA GAAAACCAGCTATCTGCAAATCAGATTTGTCTTTCACCACTTATCACAACTAATCA GATGATTTTGCAACATCAACCTGTTCGACCGTTTTATTAACTTTTTTATTTATTAA TTTTTTCCGTCTTGTCATAATAAGATCATTTGCTTTCGGGTCAATTAATATTAACT TTTTACACAATTTTTTTTTTATTCAATTTTTTTTTAATTTTTTTATTTTGCTAATA TTAATTACTTGCTGACCCATTATACAAAAGGTCCA
pIL17-RsaI-K1	TGATACGCATCMYTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCAAGAACTGTTGATTTTTC GTAACTACGTTTTTCTAGTTTTGTTTTGTTTTGTTTTGTTTTGTTTATTCCAATTT TTTtATAGTCCCATTCCTTTAGAATTTTCAACGACTATTATTTCATTTTTATAAAA

	CATATTTGTTGTGTCAATGCCACAAAAGTAAAAGGCGTCACAAAAAACAAAAAAAA AGAAGAGTAAAACAATCAAGTAGCCTATTTAATATGGATAGAAACAAACAAGTCTT GCTTCAAGTATAACCTAAGATTAATTTCTTTTGATTGCACCCAACATTTTACAAGG TATGTCCGACGACGTTCAAGAGGCCGAAACTGTCGCATTAGTCAATGACAAAATTT ATAATTCCTCAACTCATTCTCCAAAAAAGACTGTCCCAAAAAAAGGTAATTTTACA TTTCAAGAGTCAAATAAGGACATATGCCTGAAAAACTATGAGGGCGAAACTACTCA AAATGAGAGGAGTGAAATTCTTCGGCAATCCATCAATCATAAGCAAGCTTCTCCTC TGAAAGTAGACACCGATAAGAAGGGCATAGAGCCTGGCATGAGCTACTTTCTAGCA CCTTCGCCAGGCTTGCAAATGAGCTTAGAATCCGATAACTTTGTGGACAGACCGCC GCTTAAAAAaGGATTACATCATTGGTATCGATTCAATACCACAGCAAAGAGAAGCG TTGATCGCTCACAAAGGTGGACACCTCACTATCATGCTTGCTGGACAGTCTGGACT AGGTAAAACAAaCttTTKKGRAMMCGTTGTTKGGtCCA
pIL17-RsaI-K4	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTTAAGAACAAAGCAAGAG AAAAAAACAAAAAAAAATCACGATATTAAATAGTGATAATACTGTGCATATGTAAA TCGTTATAAAGCATATAGTTTACGCTCTCGAATGTTAAGCAACACTTGTTATTAAG CCTTGTTGGCGCAATCGGTAGCGCGTATGACTCTTAATCATAAGGCTAGGGGTTCG AGCCCCCTACAGGGCTTTTATTTTTTCGTATCATCAGCTAAACGCTGAAGAATCTG CTCATTTTTTCGATAGGCGAAACACACAAAGAATTAAAAAAAAAAGATTAAACGAA AAATAGTCTAGTTAGAATTTTATCAGACGATATAATCATATTTTGGGTGTCCA
pIL17-RsaI-K6	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGACTTTTCTTGCGCTTTTAA ACTTTCCATCGCAAGTGAACATATAAAATCTGAAAATGACAACAGACATTGTAAAC AAATAAAATGGATTATGGTTCATTCCCTT
pIL17-RsaI-K7	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCCACTATATACCATAGTCT TTGTAGTTTTAGTTAATTGTTCCGAATTTGAATTTCTTAAATTATGTTTAAAGTAT CAATCGGCTTCAATAACATGTAGCAAATACGCTTTTAAAAAGAAAAAAAACCTTTT AGAGTCAAATATTATCGATGGAAAATACTTAAaGATTGATACACACATCAGCTGTA AACTCAGTTTCACAAGTAAACTTCAAACAGTTTGTCCCATCAGGAAACCCTCCTTC CTGATTGTTATGCTCTTTTGCTATCACATTGCACAAAGAATTCGAGTAAAAAGGTT TAATCCCTTGCTGCAAAATACCAATAGTTATGTGGGGAGAATCGTTGGCACATTCT AGTCTTGAAACTTTGAGTTCATTGGTGGAGTCATACACACATTCCTTCTCAGACTT GACCACAATTGTAACGATCTTGTCATCCCAGAGTAACTTGTCTAGTTTGAATTTCA TCACATCTTTCGTCTTGATTAGTTTTGGCCA
pIL17-RsaI-K8	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGACGCAAGACTTGTCGGAAA TAAAATGGTTAGTATCGTAAGGACAAGCGAATTCAAGGAAAACAAACATATTGCTA TTATTACAGAGGGTAGATAGGACGAGGTGTCTTAAGCCTTTTCTAGTGCATCAAAT ATATATTTTATGGAGATACATGATACATATTCCTAAACATTGTAAGGCCAAGGGAT ACCTCTTACTTATTTTTATTATTTTCGTTGGCGTGTTTTGTTTTTAGAATTATGAG AGCAGTGAACTTTATAAAAAAATAAAGCCTTCTGAAAATATAAACAAAATAAATAA ATGATAAAAAAATGTGAAAACCACGAAAACAAAACAACGATTAAAAGTTCTATGGT TGATTTTTTTTTGGTCCACGGCAAACTTACATTAGATCTATCTCTTTTTTCTTTTT TTTTTTTTTTTTCCAGCGAATATGTTGATTAGCAGTTTGACTTTAAATTTGCCA
pIL17-RsaI-K9	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGACCGTTCTTTAATAAATTT ATATTATTTTTATTAATAATTATATTTATAAAGGCTATACTTACATATATTTCATA GAAAACCAGCTATCTGCAAATCAGATTTGTCTTTCACCACTTATCACAACTAATCA GATGATTTTGCAACATCAACCTGTTCGACCGTTTTATTAACTTTTTTATTTATTAA TTTTTTCCGTCTTGTCATAATAAGATCATTTGCTTTCGGGTCAATTAATATTAACT TTTTACACAATTTTTTTTTTATTCAATTTTTTTTTTAATTTTTTTATTTTGCTAATA TTAATTAC TTGCTGACCCATTATACAAAAGGCCA
pIL17-RsaI-K10	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGCTTAAATGAAAATAAATAA TGATGGCGGTAATGATAAAAGGTATAATAATAAAATAATAATAAATAATTGTTTGT

	TCTAGCAACTAAGAATCAATCGTTAGTCATGAAATCAATAAATTTACAACAAGTTT TAGTCTGTTGCCAAGGCCTTTTTAATTTCGGCAATAGCCTTACCTGGATTCAAACC CTTTGGACAGGTTCTGGTCCA
pIL17-RsaI-K11	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGATCGAGCAGAGAAAGCTGT TGGTCGACGTTCAGGCCAAGGCCGAAGACGAAATGGAAGTGGACGATGCCAGTGCC GTTGTGCTAACAAAGAAACAAAAGCCAGTGGACCCTTTTCAATTGTTGATCTTGAC ACTCATTTCTAGATTTTTTTTTGTTTTTCAAGGTTTTTTTTTATTTTAGGTTGCAT AAAGAATAAGAAAACTAAAGTCTCGTAACATAAAATATAA ATTTCAAAAATTTAGTCCA
pIL17-RsaI-K12	TGATACGCATCACTATGCAATTGTTAGGGCTCCCATATATTTAGGTTCAGTAGTTG CTCTGTATGTTTAATGTTTTTACGCTGGAAGGTCTGGACGGTAGCGACTAAAATAG CACTTTTGCACCCCTGGGTTTTTTCCACAAACCAAAAAAAGCGTCAAAAGCTAGAT AGTTACAAAACAATTCCCAATTTGGTATCGTAATAATAAGTCTTAATATTTGCCCG TAAATAGTTGAGATTGACATATGAAAACCACTGTTTGCTTTGCATACTCTGGCTTT TTAGAAATGTAGGGCGGGTATGACTTGTGTGTTTTCTTTTTTCTTGAATGGTTTTA GTATATAAATGTCGGAGATTGAAGTAAATTCACACAACCTTGATGTCCTTACCCGA CTTATTTGATGAAATGATCACTTAGTTGCGTATACGATCGGTTTTAATGCTATGCT CCTTTCTTCCTTAGTCTTTAGTTTTTTTTTCTTGGAAAATAATATTGTAATGTAAT TCATGTTGTTTCAGGAGAAATTGGATAAAAAAAAGATTCAGATTAAACAACTTGAA AAAAAATCGATATACCGCATGTGTAAGCAGATAGACCGAACCAAACCCCTGTCCA

Table 5: Sequences of ARSs and flanking elements isolated using pIL17*. The inserts must be tested in pRS406 to determine whether they are ARSs or flanking elements.

*Sequence not yet available for all isolated inserts.